Supplemental Material 1

Data collection

1. *The Anxiety Disorders Interview Schedule for DSM-IV (ADIS)*. The ADIS [1] was administered by masked evaluators. They received a one-day training by the 3rd author (MS) who was experienced clinical psychologist regarding implementing the ADIS and observed previous interview sessions. Diagnoses were based on information provided by both informants (i.e., composite diagnoses). First, a child and parent participate in the interview in the same room, then parent only interview is applied. For child and parent, Separation Anxiety Disorder, Social Anxiety Disorder, Specific Phobia, Generalized Anxiety Disorder, Persistent Depressive Disorder (dysthymia), Depression, Obsessive Compulsive Disorder, Panic Disorder, and exclusion criteria are evaluated. For parent only interview, Attention Deficit/Hyperactivity Disorder, Oppositional Defiant Disorder, Intellectual Disabilities, Specific Learning Disorder, Autism Spectrum Disorder, School Absenteeism, and Selective Mutism are evaluated. Primary diagnoses were determined based on the severity of disorders (Clinical Significance Ratings: CSR) as well as judgements by the evaluator.

2. *Spence Children's Anxiety Scale (SCAS) for child and parent*. The SCAS [2] is a 38-item self-report measure of anxiety symptoms designed for children and adolescents. The Spence Children's Anxiety Scale-Parent version (SCAS-P) [3] is parent version of the SCAS and also consists of a 38-item. Each item is rated on a 4-point scale in terms of its frequency ranging

from 0 (never) to 3 (always). The total score of all 38 items represents overall anxiety score. Ishikawa and colleagues developed the Japanese version of the SCAS [4] with strong internal reliability coefficients: .94 and .92 for the full-scale scores among children and adolescents, respectively. In addition, the scale has acceptable test–retest reliabilities of 2-4 weeks: $r = .76$ for children and $r = .86$ for adolescents, $ps < .001$ [4]. The Japanese version of the SCAS-P was also developed [5]. Internal reliabilities of the Japanese version were satisfactory for the community and clinical samples (Cronbach's alpha with corrected Spearman Brown coefficients = .96).

3. *Depression Self-Rating Scale (DSRS)*. The DSRS [6] is an 18-item measure of depressive symptoms for children and adolescents between 6 and 15 years of age. The total score of the DSRS represents depressive symptoms and each item is rated on a 3-point scale in terms of its frequency from 0 (never) to 2 (always). Murata (1996) translated the English scale into Japanese (cited by 7). The test-retest reliability of Japanese scale was .73 and Cronbach' alpha was .77.

4. *Children's Cognitive Error Scale (CCES)*. The CCES [8] is a 20-item measure to capture cognitive errors in children and adolescents. Each item is rated on a 4-point scale in terms of its frequency from 0 (never think so) to 3 (think so very much). Explanatory and

confirmatory factor analyses revealed that the CCES has a single factor structure [8, 9]. Two-week test-retest reliability of the CCES was r=.66, p< .01 and Cronbach's alpha was .85 [9].

5. *Family Accommodation Scale for Anxiety (FASA) for parent and child.* FASA is a parent-rated scale including 16 items rated on a 5-point Likert-type scale ranging from 0 to 4 [10]. FASA yields an overall Accommodation score (9 items), and subscale scores for Participation (4 items), Modification (5 items), Distress (1 item), and Consequences (3 items). Child reported FASA (FASA-CR) also consists of 16 items and consistent subscales with FASA [11].

6. *Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT).* The CAS-CBT is an 11-item scale developed with the purpose of measuring adherence and competence in CBT for anxiety disorders in children and adolescents [12]. The CAS-CBT is composed of two major domains Adherence and Competence. Adherence was rated on a 7-point scale from 0 (None) to 6 (Thorough) for seven items and Competence was rated on a 7-point scale from 0 (Poor skills) to 6 (Excellent skills) for four items evaluated by CBT therapists/supervisors. This study uses the Japanese version which was translated by the first author with permission of the first author of the original study. In this study, supervisors who are in charge of each site (i.e., Kyoto, Hyogo, Nagano) evaluate each therapist independently.

7. *Cross-cultural Behavioural Observation System (C-BOS).* The C-BOS [13] was developed for cross-cultural comparison of CBT programs delivered to children and adolescents of different cultures. The C-BOS is designed to evaluate four major domains: readiness, dynamics, accommodation, and orientation. Independent raters evaluate the C-BOS based on the recorded video files for each session. The detailed evaluation system of the C-BOS is described the previous studies [14].

8. *Satisfaction and comprehension*. The original scale is developed for the purpose of this study consisting of 9 items for child and 14 items for parent. The questions ask to rate on a 4-point scale how useful the program and to what extent it is satisfiable. In addition, an open-ended question for both child and parent is prepared to describe any feedback from participants.

References

1. Silverman, W. K., & Albano, A. M. Anxiety disorders interview schedule for DSM-IV: Child version. Graywind Publications Inc 1996.

2. Spence, S. H. A measure of anxiety symptoms among children. Behav Res Ther 1998; 36(5), 545-566.

3. Nauta, M. H., Scholing, A., Rapee, R. M. et al. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. Behav Res Ther 2004; 42(7), 813-839.

4. Ishikawa, S., Sato, H., & Sasagawa, S. Anxiety disorder symptoms in Japanese children and adolescents. J Anxiety Disord 2009; 23(1): 104–111.

5. Ishikawa, S., Shimotsu, S., Ono, T. et al. A parental report of children's anxiety symptoms in Japan. Child Psychiatry Hum Dev 2014; 45(3):306–317.

6. Birleson, P. The validity of depressive disorder in childhood and the development of a self-rating scale: a research report. J Child Psychol Psychiatry 1981, 22(1), 73-88.

7. Murata, T., Shimizu, A., Mori, Y., & Oushima, S. Childhood depressive state in the school situation: Consideration from the Birleson's scale. Jpn J Psychiatry 1996; 1: 131-138.

8. Ishikawa, S. Cognitive errors, anxiety, and depression in Japanese children and adolescents. Int J Cogn Ther 2012; 5(1): 38-49.

9. Ishikawa, S., & Sakano, Y. Cognitive error and trait anxiety in children: Development of a children's cognitive error scale. Jpn J Behav Ther 2003; 29: 145-157.

10. Lebowitz, R. E., Woolston, J., Bar-Haim,Y. et al. Family accommodation in pediatric anxiety disorders, Depress Anxiety 2013; 30 (1): 47-54.

11. Lebowitz, R. E., Scharfstein, L., & Jones, J. Child-report of family accommodation in pediatric anxiety disorders: Comparison and integration with mother-report. Child Psychiatry Hum Dev 2015; 46(4), 501-511.

12. Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W. et al. Competence and adherence scale for cognitive behavioral therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. Psychol Assess 2016; 28(8): 908.

13. Ishikawa, S., & Hudson, J. L. Cross-cultural behavioural observation system: Coding behavioural coding manual and calculation sheets (3rd ver.). Faculty of Psychology, Doshisha University 2019.

14. I shikawa, S., Romano, M., & Hudson, J. H. A comparison of interactions among children, parents, and therapists in cognitive behavior therapy for anxiety disorders in Australia and Japan. Behav Ther 2022; 53(1): 34-48. https://doi.org/10.1016/j.beth.2021.05.008